Accepted Manuscript

Title: Can we predict detrusor overactivity in women with lower urinary tract symptoms? The King's Detrusor Overactivity Score (KiDOS)

Author: Ilias Giarenis Patrick Musonda Heleni Mastoroudes Dudley Robinson Linda Cardozo

PII:	S0301-2115(16)30822-3
DOI:	http://dx.doi.org/doi:10.1016/j.ejogrb.2016.07.495
Reference:	EURO 9523
To appear in:	*EURO*
Received date:	3-2-2016
Revised date:	18-6-2016
Accepted date:	26-7-2016

Please cite this article as: Giarenis I, Musonda P, Mastoroudes H, Robinson D, Cardozo L, Can we predict detrusor overactivity in women with lower urinary tract symptoms? The King's Detrusor Overactivity Score (KiDOS), *European Journal of Obstetrics and Gynecology and Reproductive Biology* (2016), http://dx.doi.org/10.1016/j.ejogrb.2016.07.495

Title

Can we predict detrusor overactivity in women with lower urinary tract symptoms? The King's Detrusor Overactivity Score (KiDOS)

Authors

Ilias Giarenis[1], Patrick Musonda[2 3], Heleni Mastoroudes[1], Dudley Robinson[1], Linda Cardozo[1]

[1] *Department of Urogynaecology, King's College Hospital, London, UK*

[2] *School of Medicine, Health policy & Practice, University of East Anglia, Norwich, UK*

[3]*Centre for Infectious Diseases Research In Zambia (CIDRZ), Plot 1275 Lubuto Road, PO Box 34681, Lusaka, Zambia.*

Corresponding author

Ilias Giarenis

Department of Urogynaecology, Norfolk and Norwich University Hospital, Norwich, NR4 7UY, UK

Telephone: 0044 1603 287102

e-mail: ilias.giarenis@nhs.net

Condensation

A prediction model is able to predict detrusor overactivity more accurately than a symptomatic diagnosis alone, in women with lower urinary tract symptoms.

Abstract

Objective

Traditionally, urodynamic studies (UDS) have been used to assess lower urinary tract symptoms (LUTS), but their routine use is now discouraged. While urodynamic stress incontinence is strongly associated with the symptom of stress urinary incontinence (SUI) and a positive cough test, there is a weak relationship between symptoms of overactive bladder and detrusor overactivity (DO). The aim of our study was to develop a model to predict DO in women with LUTS.

Study Design

This prospective study included consecutive women with LUTS attending a urodynamic clinic. All women underwent a comprehensive clinical and urodynamic assessment. The effect of each variable on the odds of DO was estimated both by univariate analysis and adjusted analysis using logistic regression.

Results

1006 women with LUTS were included in the study with 374 patients (37%) diagnosed with DO. The factors considered to be the best predictors of DO were urgency urinary incontinence, urge rating/void and parity (p-value <0.01). The absence of SUI, vaginal bulging and previous continence surgery were also good predictors of DO (p-value < 0.01). We have created a prediction model for DO based on our best predictors. In our scoring system, presence of UUI scores 5; mean urge rating/void ≥ 3 scores 3; parity ≥ 2 scores 2; previous continence surgery scores -1; presence of SUI scores -1; and the complaint of vaginal bulging scores -1. If a criterion is absent, then the score is 0 and the total score can vary from a value of -3 to +10. The Receiver Operating Characteristic (ROC) analysis for the overall cut-off points revealed an area under the curve of 0.748 (95%CI 0.741, 0.755).

Conclusion

This model is able to predict DO more accurately than a symptomatic diagnosis alone, in women with LUTS. The introduction of this scoring system as a screening tool into clinical practice may reduce the need for expensive and invasive tests to diagnose DO, but cannot replace UDS completely.

Keywords

Detrusor overactivity, lower urinary tract symptoms, prediction model, urinary incontinence, urodynamics

Text

Introduction

Lower urinary tract symptoms (LUTS) are common in women with a reported prevalence up to 66.6% in large epidemiological studies [1]. They are under-reported and under-treated despite their significant adverse effect on quality of life (QoL) [2]. Traditionally, urodynamic studies (UDS) have been used to assess LUTS, but their routine use is now questioned. As UDS are invasive and expensive tests without evidence-based additional value in the management of women with urinary incontinence (UI), their routine use is discouraged by international professional bodies [3,4].

Conventionally, stress urinary incontinence (SUI) is the predominant symptom associated with urodynamic stress incontinence (USI) and urgency urinary incontinence (UUI) with the urodynamic observation of detrusor overactivity (DO). The correlation between the clinical and urodynamic diagnosis in women is weaker than the correlation in the male population [5]. Several studies have compared clinical and urodynamic diagnoses of UI in women. A recent systematic review showed that the positive predictive value (PPV) of SUI to diagnose USI was 75% (range 41-95%), but the PPV of UUI to diagnose DO was only 58% (range 22-100%) [6]. A combined assessment including history, physical examination and bedside tests is less helpful in diagnosing DO compared to USI [7].

Few researchers have attempted to identify clinical predictors of DO. Variables such as the maximum urge rating in an urgency scale [8], the presenting bladder volume at urodynamics [9] have been proposed as potential predictors. A number of prediction tools have been developed to help to diagnose DO, but their use in clinical practice has been limited due to the required complex calculations [10,11]. A prediction model combining symptoms, examination findings and non-invasive tests such as bladder diaries incorporating urgency scales is probably closer to daily clinical practice and is likely to show better agreement with UDS. The purpose of this study was to develop a simple prediction model to estimate the risk of DO in female patients with LUTS.

Materials and methods

Participants

This was a cross-sectional study in a tertiary referral Urogynaecology Unit. Consecutive women attending a one-stop urodynamic assessment clinic with LUTS were included in the study. Ethical approval was granted by the regional Research Ethics Committee.

We excluded women unable to read and complete a questionnaire in the English language; younger than 18 years; with dementia or memory disorders; with known neurological conditions such as stroke, multiple sclerosis, spinal cord injury or Parkinson's disease; on antimuscarinic medication within seven days of the attendance in the clinic; with evidence of urinary tract infection on urinalysis (presence of nitrites with or without leucocytes) on the day of the appointment.

Procedures

All women were asked to complete a disease-specific health related quality of life (HRQoL) questionnaire (King's Health Questionnaire, KHQ) [2] and 3-day bladder diary incorporating the validated Patient's Perception of Intensity of Urgency Scale (PPIUS) [12], before attending the urodynamic clinic. The symptom domain of the KHQ was used to assess the presence of LUTS. The five grades of the PPIUS (from 0: no urgency, to 4: urgency incontinence) were used to assess the degree of urgency associated with each void. Urgency episodes were counted as suggested by Cardozo et al as voids with PPIUS level 3 and 4 (without or with urgency incontinence respectively) [13]. Daytime urinary frequency, nocturnal frequency and the functional bladder capacity were recorded from the bladder diary.

Initial assessment included medical history, physical examination and urinalysis. Pelvic organ prolapse was assessed in both the lithotomy position and standing with the patient exerting a maximal Valsalva manoeuvre using the pelvic organ prolapse quantification (POPQ) system [14]. The participants then underwent multichannel urodynamics according to the ICS recommendations [15]. Women, whose symptoms of urgency were not reproduced during the laboratory test, underwent a 4-hour ambulatory urodynamics test following a standardised protocol [16].

<u>Statistical analysis</u>

Analysis of descriptive data was carried out in three ways. Firstly we looked at continuous variables and investigated their distribution. If a variable was found to be symmetrical, then the mean and standard deviation was used to summarise the variable, otherwise the median and interquartile range was used to summarise the variable. To test any differences for continuous variables either a parametric or non-parametric test was used depending on whether the variable was symmetrical or not. Secondly, we considered nominal categorical variables and tabulated the proportions in the DO group and non-DO group and the differences were tested by using either the chi-squared test or Fisher's exact test depending on whether the expected assumptions for chi-squared test were satisfied or not. Thirdly, we considered categorical variables with ordinal values. For these variables, proportions were analysed and the Cuzick's test for trend was carried out [17].

The effect of each of the 27 variables on the odds of DO was estimated both by univariate analysis and adjusted analysis using logistic regression. Assuming a bladder diary completion rate of 75% and a DO prevalence of 36% in our population [18] we estimated a minimum sample size of 1000 patients based on the work by Peduzzi et al (N = 10 k / p, where k equals the number of covariates and p the smallest of the proportions of negative or positive cases in the population) [19].

An investigator-led best model selection approach was used to select the best predictors of DO in the multiple logistic regression model, as opposed to machine-led step-wise regression, which is not advisable. For ordinal explanatory variables, Mantal-Haenszel odds ratio for trend was used to estimate the odds ratio taking account of the ordinal nature of the data [20]. To deal with missing data in the adjusted model, we explored complete case analysis but also used multiple imputation methods [21,22]. We carried out 20 multiple imputations and estimated the best predictors of DO. We used estimates from the adjusted model after multiple imputations to create the prediction scoring tool. To test how the prediction tool agreed with the observed values, Kappa statistics was calculated. The overall predictive ability for our prediction model was measured by the area under the Receiver Operating Characteristic (ROC) curve. Sensitivity, specificity and likelihood ratio for each cut-off point of the scoring system were calculated. All analyses were performed using STATA software, version 12.1 SE (Stata Corporation, College Station, TX, USA).

Results

1006 women with LUTS were included in the study. The mean age was 51.4 years (SD: 14.8) and 52 % of them were postmenopausal. 374 patients (37%) were diagnosed with DO. The basic characteristics of our population with the univariate comparison are presented in table 1.

As a result of the multiple logistic regression with multiple imputations the investigating team determined that the factors considered best predictors of DO were UUI, urge rating/void and parity. The absence of SUI, no vaginal bulging and no history of previous continence surgery were also good predictors of DO (table 2).

Using estimates from the adjusted model and the relevant odds ratios (table 2), we developed a prediction model called King's DO Score (KiDOS). The scoring system is such that for those that had decreased effect in DO, a negative score was given, for those with an increased effect, a positive score was given with respect to odds ratio. In KiDOS, presence of UUI scores 5; mean urge rating/void ≥ 3 scores 3; parity ≥ 2 scores 2; history of previous continence surgery scores -1; presence of SUI scores -1; and the complaint of vaginal bulging scores -1 (table 3). If a criterion is absent, then the score is 0 and the total score can vary from a value of -3 to +10.

The overall predictive ability for our prediction model, measured by the area under the Receiver Operating Characteristic (ROC) curve was 0.748 (95%CI 0.741, 0.755) (figure 1). The agreement between the predicted and the observed values was found to be kappa=0.685 (95% CI 0.597, 0.721).

The overall sensitivity, specificity and likelihood ratio for each cut-off point are shown in table 4. The statistical optimal cut-off point for our score was 6 (table 5), achieving a sensitivity of 62.4% (95%CI 61.3%, 63.6%), a specificity of 72.2% (95%CI 71.4%, 73.0%), a positive predictive value of 57.9% (95% CI 56.7%, 59.0%) and a negative predictive value of 75.9% (96% CI 75.0%, 76.7%). In comparison, in our cohort the symptom of urgency had a sensitivity of 96.2% and a specificity of 28.1% to predict DO and urgency incontinence a sensitivity of 88.2% and a specificity of 47.6% (table 6).

Comment

As the routine use of UDS is currently discouraged, alternative methods have been explored in order to improve the clinical prediction of DO. This study describes the development of a prediction model with an associated scoring system for DO (KiDOS). This model, combining symptoms, medical history and non-invasive tests, is able to predict DO more accurately than a symptomatic diagnosis alone, in women with LUTS.

The KiDOS is an easily calculated scoring system that includes six variables obtained during a standard assessment of LUTS. The predictors can be established with history taking (parity, previous continence surgery) and completion of a validated symptom questionnaire (UUI, SUI, vaginal bulging) and a bladder diary (urge rating/void). Our prediction model was developed using a large sample with a prior power calculation and following robust statistical methods. The diagnostic accuracy of our prediction model based on the ROC analysis will be classified as "fair" (0.7-0.8) with an area under the ROC curve of 0.748. The value of Kappa statistics (0.685) correspond to "good" agreement between the predicted values and the observed urodynamic findings [23]. Based on the performance of the prediction model we could argue that it can improve symptomatic diagnosis but cannot replace UDS altogether.

A number of authors have attempted to correlate symptoms, examination findings and frequency-volume variables with a urodynamic diagnosis of DO [10,11,24,25,26]. Contreras Ortiz et al described the Bladder Instability Discriminant Index (BIDI), a clinical index derived by 7-day frequency-volume charts from only 89 patients [10]. Despite a reported sensitivity of 82% and specificity of 87% its use in clinical practice has been limited due to the required complex calculations. Some of these variables (daytime / nighttime voids and functional capacity) were included as potential predictors in our analysis. In the univariate comparison, nighttime voids (p=0.001) and functional capacity (p<0.001) but no daytime voids (p=0.143) were statistically significant. However, their potential contribution to the model following the adjusted analysis with logistic regression was limited and therefore not included. In our study the bladder diary included the validated PPIUS and the strongest predictor was mean urge rating/void. This finding reinforces the current knowledge that urinary urgency, rather than nocturia or daytime frequency, is the defining and cornerstone symptom of OAB and DO [27].

Regarding LUTS it has been shown that responses to self-completed postal questionnaires have a better relationship with urodynamic findings compared to interview-assisted questionnaire responses [24]. Matharu et al also assessed the relationship between LUTS reported in a self-completed questionnaire and urodynamic diagnosis [25]. In their multivariate model for DO they included UUI, strength of urgency, and nocturia as positive predictors and SUI as a negative predictor. Their risk model had a sensitivity of 63% and a specificity of 65%, but the results of a ROC analysis were not presented. In line with this study we found that UUI was the strongest positive predictor (aOR 4.10) and SUI the strongest negative symptom predictor (aOR 0.45) for DO. In our study the subjective strength of urgency and symptom of nocturia were expressed as objectively measured urge rating/void and nighttime voids on the bladder diary.

A study of 171 women by Vella et al attempted to develop a scoring system combining demographic data, LUTS and frequency-volume chart data [11]. The nine variables (with the contributing score) included in the final questionnaire were: urgency (11), voided volume < 300ml (9), UUI (8), nocturia (6), nocturnal enuresis (6), flooding (6), nulliparity (5), frequency (2), non-caucasian ethnicity (2). In view of the use of multiple variables in a small sample without adjusted logistic regression analysis, there is a significant risk of overfitting for this scoring system. In our model parity ≥ 2 rather than nulliparity was a predictor of DO and our findings are consistent with epidemiological research relating to OAB [28].

A recent study, published after the international presentation of our study, used a similar design in a cohort of 1140 women [26]. In their final logistic model they included UUI, urgency and nocturia as positive predictors and SUI, parity ≥ 2 and signs of prolapse as negative predictors. This model had a sensitivity of 0.61, a specificity of 0.73 and an area under the ROC curve of 0.737. This large study, excluding the contradicting results about parity, enforces the validity of our findings in a different population sample. However, as this study did not develop a scoring system and did not report a logistic regression equation, its clinical value is limited. The difference in the statistical methods has probably led the authors of that study to consider their findings not clinically useful. Regarding the addition of POP as a negative predictor (symptom of vaginal bulging in the current study, sign of POP in the study by Haylen et al) we can postulate that POP could induce the "irritative" symptoms of urinary urgency, without inducing the pathophysiological mechanism of DO. Similar to this mechanism post-operative formation of scar tissue and localised denervation after continence surgery could explain the performance of prior continence surgery as negative predictor.

The idea behind the scoring system is to balance the scores and as such, the higher the score, the more likely the model indicates that someone has DO. The proposed cut-off point for KiDOS is 6, because this cut-off balances out the sensitivity and specificity and achieves reasonable PPV and NPV. Using the KiDOS as a screening tool at the initial assessment of all women with LUTS could help physicians to arrange conservative and drug therapy appropriately. For women with score of 5 or less, supervised pelvic floor muscle training and bladder retraining for 3 months might be offered before consideration of drug therapy (figure 2). However, depending on the setting and the use of the model in a management algorithm, different cut-offs could be selected in order to optimise sensitivity or specificity. If KiDOS is used to reduce the need for UDS after failed conservative therapy prior to stress incontinence surgery, we would consider a cut-off of 4 in order to increase the detection of DO pre-operatively.

Another approach in order to reduce the need for invasive urodynamics is the use of biomarkers as predictors of DO. Despite the initial promising results [29], a recent multi-centre study of 687 women assessing the diagnostic accuracy of BWT failed to reproduce similar findings with a reported sensitivity of 42%, specificity of 62% for BWT >5mm and an area under the ROC curve of 0.53 [30]. Near infrared spectroscopy has also been shown to be an unreliable method for detecting DO in women with OAB symptoms [31]. While several small studies have demonstrated a trend towards higher urinary nerve growth factor (NGF) in patients with DO, a recent systematic review of the literature has stated that the data are imprecise and hence NGF cannot be recommended for use in current clinical practice [32]. Compared to these biomarkers, our feasible prediction model combines an overall superior diagnostic accuracy with minimal cost as opposed to UDS which are considered to be expensive and invasive.

We acknowledge that our study had some limitations, the first being the lack of external validation of our prediction model. While our large sample and our robust statistical methods have reduced the risk of overfitting, external validation is required before introduction of the model into clinical practice. The clinical role of the model will remain questionable until external validation studies demonstrate its value. Another limitation of the study is the missing data of the bladder diary variables. The impact of those data on our calculations was regulated by multiple imputations. The routine use of urgency scales in clinical practice might decrease patient compliance in bladder diary completion and hence calculations of urge rating / void may lack accuracy. Finally, as this is a single-centre study in a tertiary referral unit the results may not be generalised for all women with LUTS or for centres that perform urodynamics following different protocols.

Diagnostic prediction models can help decision making and reduce the need for expensive investigations. Their adoption in other fields of obstetrics and gynaecology such as gynaecological oncology [33] highlights the need for similar development in urogynaecology. Sophisticated powerful instruments such as artificial neural networks might further improve the performance of risk assessment tools [34]. However, we must appreciate the fact that DO is a multifactorial condition and thus no model may be sufficiently selective and sensitive [35].

In conclusion, this model is able to predict DO more accurately than a symptomatic diagnosis alone in women with LUTS. The introduction of this easily calculated scoring system as a screening tool into clinical practice may reduce the need for expensive and invasive tests to diagnose DO, but cannot replace UDS completely.

Legends

Figure 1: Receiver Operating Characteristic (ROC) analysis

Figure 2: Proposed algorithm for management of women with lower urinary tract symptoms (LUTS)

References

1. Irwin DE, Milsom I, Hunskaar S, et al. Population-based survey of urinary incontinence, overactive bladder, and other lower urinary tract symptoms in five countries: results of the EPIC study. Eur Urol 2006;50:1306-15.

2. Kelleher CJ, Cardozo LD, Khullar V, Salvatore S. A new questionnaire to assess the quality of life of urinary incontinent women. Br J Obstet Gynaecol 1997;104:1374-79.

3. Lucas MG, Bosch RJ, Burkhard FC et al. EAU guidelines on assessment and nonsurgical management of urinary incontinence. Eur Urol 2012;62:1130-42.

4. Winters JC, Dmochowski RR, Goldman HB et al; American Urological Association; Society of Urodynamics, Female Pelvic Medicine & Urogenital Reconstruction. Urodynamic studies in adults: AUA/SUFU guideline. J Urol 2012;188:2464-72.

5. Hashim H, Abrams P. Is the bladder a reliable witness for predicting detrusor overactivity? J Urol 2006;175:191-5.

6. Van Leijsen SA, Hoogstad-van Evert JS, Mol BW et al. The correlation between clinical and urodynamic diagnosis in classifying the type of urinary incontinence in women. A systematic review of the literature. Neurourol Urodyn 2011;30:495-502.

7. Holroyd-Leduc JM, Tannenbaum C, Thorpe KE, Straus SE. What type of urinary incontinence does this woman have? JAMA 2008;299:1446-56.

8. Chung SD, Liao CH, Chen YC, Kuo HC. Urgency severity scale could predict urodynamic detrusor overactivity in patients with overactive bladder syndrome. Neurourol Urodyn 2011;30:1300-4.

9. Haylen BT, Yang V, Logan V, Husselbee S, Law M, Zhou J. Does the presenting bladder volume at urodynamics have any diagnostic relevance? Int Urogynecol J Pelvic Floor Dysfunct 2009;20:319-24.

10. Contreras Ortiz O, Pellicari A, Doretti G, Lombardo RJ. A clinical index (BIDI) based on Frequency/Volume Study in the diagnosis of bladder instability. Int Urogynecol J Pelvic Floor Dysfunct 1992;3:114-7.

11. Vella M, Robinson D, Cardozo L, Srikrishna S, Cartwright R. Predicting detrusor overactivity using a physician-based scoring system. Int Urogynecol J Pelvic Floor Dysfunct 2008;19:1223-7.

12. Cartwright R, Srikrishna S, Cardozo L, Robinson D. Validity and reliability of the patient's perception of intensity of urgency scale in overactive bladder. BJU Int 2011; 107:1612-7.

13. Cardozo L, Hessdörfer E, Milani R et al. Solifenacin in the treatment of urgency and other symptoms of overactive bladder: results from a randomized, double-blind, placebo-controlled, rising-dose trial. BJU Int 2008;102: 1120-7.

14. Bump RC, Mattiasson A, Bo K et al. The standardization of terminology of female pelvic organ prolapsed and pelvic floor dysfunction. Am J Obstet Gynecol 1996;175:10-7.

15. Schafer W, Abrams P, Liao L et al. Good urodynamic practices: uroflowmetry, filling cystometry and pressure-flow studies. Neurourol Urodyn 2002;21:261-74.

16. Anders K, Cardozo L, Ashman O, Khullar V. Morbidity after ambulatory urodynamics. Neurourol Urodyn 2002;21:461-3.

17. Cuzick J. A Wilcoxon-type test for trend. Stat Med 1985;4:87-90.

18. Digesu GA, Khullar V, Cardozo L, Salvatore S. Overactive bladder symptoms: do we need urodynamics? Neurourol Urodyn 2003;22:105-8.

19. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373-9.

20. Clayton DG, Hills M. Statistical Models in Epidemiology. Oxford: Oxford University Press, 1993.

21. Carlin JB, Li N, Greenwood P, Coffey C. Tools for analyzing multiple imputed datasets. Stata Journal 2003;3:226–44.

22. Carlin JB., Galati JC, Royston P. A new framework for managing and analyzing multiply imputed data in Stata. Stata Journal 2008;8:49–67.

23. Altman DG. Practical Statistics for Medical Research. London: Chapman and Hall, 1991.

24. Khan MS, Chaliha C, Leskova L, Khullar V. The relationship between urinary symptom questionnaires and urodynamic diagnoses: an analysis of two methods of questionnaire administration. BJOG 2004;111:468-74.

25. Matharu G, Donaldson MM, McGrother CW, Matthews RJ. Relationship between urinary symptoms reported in a postal questionnaire and urodynamic diagnosis. Neurourol Urodyn 2005;24:100-5.

26. Haylen BT, Chiu TL, Avery D, Zhou J, Law M. Improving the clinical prediction of detrusor overactivity by utilizing additional symptoms and signs to overactive bladder symptoms alone. Int Urogynecol J 2014;25:1115-20.

27. Brubaker L. Urgency: the cornerstone symptom of overactive bladder. Urology 2004 ;64(6 Suppl 1):12-6.

28. Zhang W, Song Y, He X, Huang H, Xu B, Song J. Prevalence and risk factors of overactive bladder syndrome in Fuzhou Chinese women. Neurourol Urodyn 2006;25:717-21.

29. Khullar V, Cardozo LD, Salvatore S, Hill S. Ultrasound: a noninvasive screening test for detrusor instability. Br J Obstet Gynaecol 1996;103:904-8.

30. Rachaneni S, McCooty S, Moddleton L, Brookes VI, Daniels J, Latthe P on behalf of the BUS study group. Preliminary results of BUS study. A multicentre study of the accuracy of bladder wall thickness in diagnosing detrusor overactivity in women. United Kingdom Continence Society (UKCS) annual scientific meeting 2014. London.

31. Mastoroudes H, Giarenis I, Vella M et al. Use of near infrared spectroscopy as an alternative to videourodynamics to detect detrusor overactivity in women with the overactive bladder syndrome. Urology 2012;80:547-50.

32. Rachaneni S, Arya P, Latthe P. Urinary nerve growth factor: a biomarker of detrusor overactivity? A systematic review. Int Urogynecol J 2013;24:1603-9.

33. Sayasneh A, Wynants L, Preisler J et al. Multicentre external validation of IOTA prediction models and RMI by operators with varied training. Br J Cancer 2013;108:2448-54.

34. Serati M, Salvatore S, Siesto G et al. Urinary symptoms and urodynamic findings in women with pelvic organ prolapse: is there a correlation? Results of an artificial neural network analysis. Eur Urol 2011;60:253-60.

35. Fry CH, Sahai A, Vahabi B, Kanai AJ, Birder LA. What is the role for biomarkers for lower urinary tract disorders? ICI-RS 2013. Neurourol Urodyn 2014;33:602-5.

Table 1 Basic characteristics of the population: Univariate comparison

	Detrusor overactivity		
Covariate	Yes, n=374 (37.2%)	No, n=632 (62.8%)	p-value
Age, mean (SD)	52.1 (15.8)	51.1 (14.3)	0.284[a]
BMI (kgm^{-2}), median (IQR)	28.0 (24.0, 33.0)	26.0 (23.0, 30.0)	<0.001[b]
Parity, mean (SD)	2.37 (1.64)	2.02 (1.39)	<0.001[a]
Racial group, n (%)			<0.001[c]
White	165 (53.9%)	419 (75.2%)	
Black	110 (34.0%)	78 (14.0%)	
Asian	10 (3.27%)	23 (4.13%)	
Mixed	3 (0.98%)	8 (1.44%)	
Other	18 (5.9%)	29 (5.21%)	
Menopause, yes, n (%)	194 (51.9%)	328 (51.9%)	0.993[c]
Hormone replacement therapy, n (%)			0.039[c]
No	350 (93.6%)	564 (89.2%)	
Local oestrogens	9 (2.41%)	35 (5.54%)	
Systemic HRTs	15 (4.01%)	33 (5.22%)	
Hysterectomy, yes, n (%)	80 (21.4%)	167 (26.4%)	0.073[c]
Prolapse surgery, yes, n (%)	31 (8.29%)	93 (14.7%)	0.003[c]

Continence surgery, yes, n (%)	36 (9.63%)	104 (16.5%)	0.002[c]
Stress incontinence, yes, n (%)	245 (65.5%)	404 (63.9%)	0.612[c]
Urinary urgency, yes, n (%)	360 (96.3%)	454 (71.8%)	<0.001[c]
Urgency urinary incontinence, yes, n (%)	330 (88.2%)	331 (52.4%)	<0.001[c]
Vaginal bulging, yes, n (%)	65 (17.4%)	155 (24.5%)	0.008[c]
Bladder pain, yes, n (%)	8 (2.14%)	21 (3.32%)	0.278[c]
Recurrent UTIs, yes, n (%)	49 (13.1%)	102 (16.1%)	0.192[c]
Voiding difficulties, yes, n (%)	53 (14.2%)	106 (16.8%)	0.274[c]
Continuous incontinence, yes, n (%)	4 (1.07%)	3 (0.47%)	0.434[d]
Nocturnal enuresis, yes, n (%)	24 (6.42%)	9 (1.42)	<0.001[c]
Day time voids, median (IQR)	8.00 (6.00, 10.0)	8.00 (6.00, 9.00)	0.143[b]
Night time voids median (IQR)	1.00 (0, 2.00)	1.00 (0, 1.00)	0.001[b]
Functional capacity, mean (SD)	379.1 (150.5)	470.0 (191.0)	<0.001[a]
Max PPIUSS, n (%)			<0.001[e]
0	0 (0.00%)	4 (1.47%)	
1	1 (0.7%)	17 (6.25%)	
2	12 (8.63%)	62 (22.8%)	

3	21 (15.1%)	102 (37.5%)	
4	105 (75.5%)	87 (32.0%)	
Mean urge rating/void, mean (SD)	2.53 (0.67)	1.81 (0.73)	<0.001[a]
Daily urgency episodes, median (IQR)	4.66 (2.67, 6.66)	1.33 (0.00, 3.66)	<0.001[b]
Voided volume (mls), median (IQR)	127.0 (70.0, 249.0)	178.5 (93.0, 346.0)	<0.001[b]
Peak flow rate (ml/sec), median (IQR)	18.0 (12.0, 29.0)	20.0 (12.0, 29.0)	0.455[b]
POPQ ordinal stage, n (%)			0.787[e]
1	244 (66.0%)	411 (65.7%)	
2	105 (28.4%)	169 (27.0%)	
3	21 (5.68%)	45 (7.2%)	
4	0 (0%)	1 (0.16%)	

BMI: body mass index, HRT: hormone replacement therapy, IQR: interquartile range, POPQ: pelvic organ prolapse quantification system, PPIUS: patient's perception of intensity of urgency scale, SD: standard deviation, UTI: urinary tract infection

[a] Two-sample t test (unequal variance)

[b] Two-sample Wilcoxon rank sum test (Mann-Whitney test)

[c] Chi squared test

[d] Fisher's exact test

[e] Cuzick's non-parametric test for trend

Table 2 Best predictors following multiple logistic regression with multiple imputations

Covariate	Adjusted odds ratio (95% confidence intervals)	p-value
Urgency urinary incontinence	4.10 (2.61, 6.48)	<0.001
Urge rating / void	2.82 (1.99, 4.00)	<0.001
Parity	1.21 (1.08, 1.36)	0.001
Continence surgery	0.38 (0.23, 0.65)	<0.001
Stress urinary incontinence	0.45 (0.31, 0.66)	<0.001
Vaginal bulging	0.57 (0.38, 0.85)	0.006

Table 3 King's Detrusor Overactivity Score (KiDOS)

	Yes	No
Urgency urinary incontinence	5	0
Urge rating / void ≥ 3	3	0
Parity ≥ 2	2	0
Continence surgery	-1	0
Stress urinary incontinence	-1	0
Vaginal bulging	-1	0
Total score (range)	(-3, +10)	

Table 4 Overall sensitivity, specificity and likelihood ratio for each KiDOS cut-off point

Cut-point	Sensitivity	Specificity	Correctly classified	LR+	LR-	d_OR
≥ -2	100.00%	0.00%	37.95%	1.000	-	-
≥ -1	100.00%	1.66%	38.98%	1.016	0.000	∞
≥ 0	99.72%	8.10%	42.87%	1.085	0.035	31.0
≥ 1	95.10%	24.18%	51.09%	1.254	0.203	6.18
≥ 2	91.37%	42.05%	60.76%	1.577	0.205	7.69
≥ 3	89.04%	46.94%	62.92%	1.678	0.234	7.17
≥ 4	87.75%	49.15%	63.79%	1.726	0.249	6.93
≥ 5	78.51%	59.33%	66.61%	1.930	0.362	5.33
≥ 6	62.43%	72.21%	68.50%	2.247	0.520	4.32
≥ 7	33.71%	89.38%	68.26%	3.175	0.742	4.28
≥ 8	23.42%	94.68%	67.64%	4.401	0.809	5.44
≥ 9	15.34%	97.11%	66.08%	5.312	0.872	6.09
≥ 10	2.61%	99.65%	62.83%	7.574	0.977	7.75
> 10	0.00%	100.00%	62.05%	-	1.000	-

$\mathrm{LR}(+) = \text{Likelihood ratio}(+\mathrm{ve}) = \Pr(+\mathrm{ve}|+\mathrm{ve}) / \Pr(+\mathrm{ve}|-\mathrm{ve})$,

$\mathrm{LR}(-) = \text{Likelihood ratio}(-\mathrm{ve}) = \Pr(-\mathrm{ve}|+\mathrm{ve}) / \Pr(-\mathrm{ve}|-\mathrm{ve})$,

$\text{d-OR} = \text{diagnostic odds ratio} = \mathrm{LR}(+) / \mathrm{LR}(-)$

∞ =Undefined

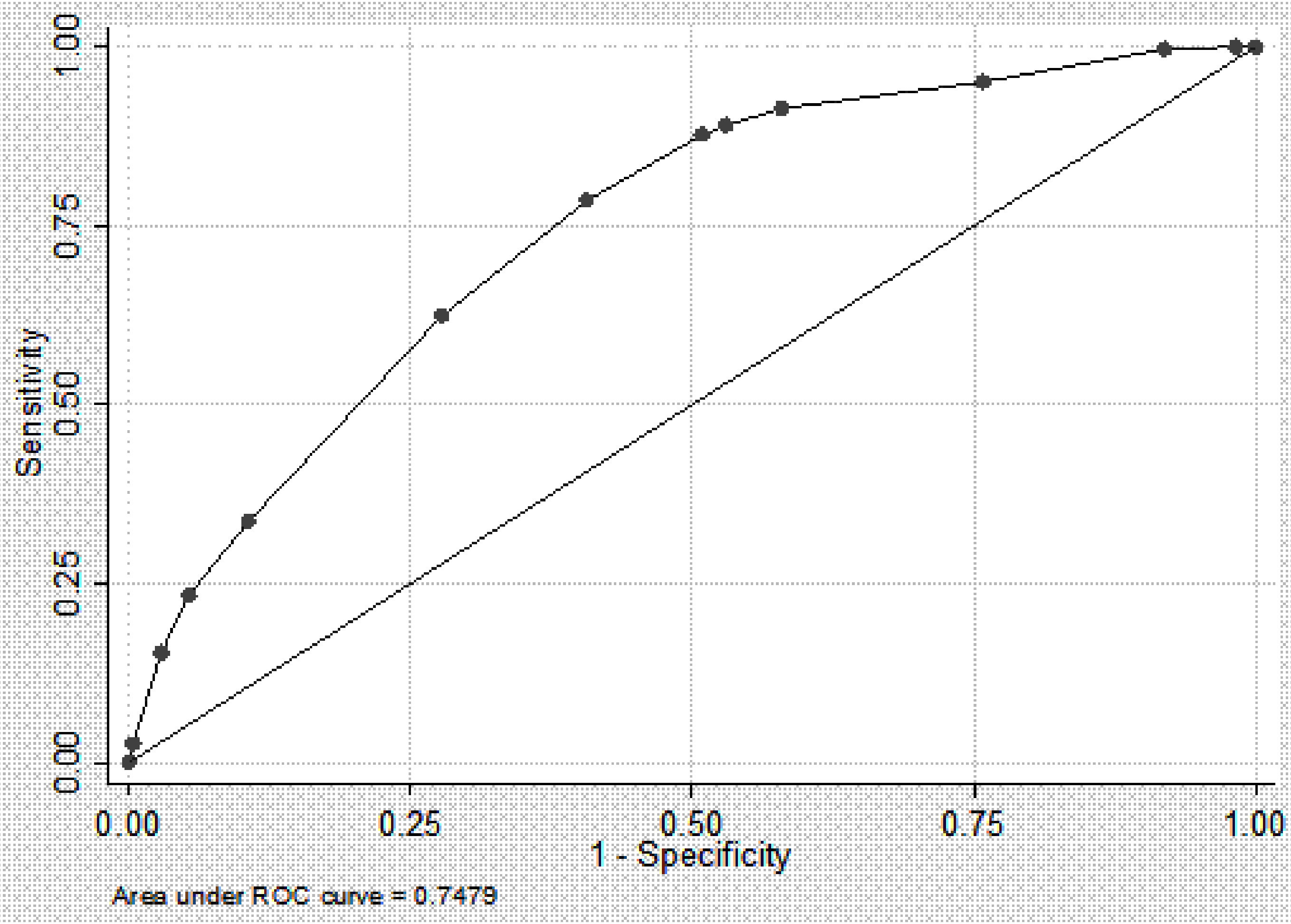
Table 5 Sensitivity, specificity, positive and negative predictive value for KiDOS cut-off of ≥6

	Estimate (95% confidence intervals)
Sensitivity	62.4% (61.3%, 63.6%)
Specificity	72.2% (71.4%, 73.0%)
Area under ROC curve	0.673 (0.666, 0.680)
Positive predictive value (PPV)	57.9% (56.7%, 59.0%)
Negative predictive value (NPV)	75.9% (75.0%, 76.7%)

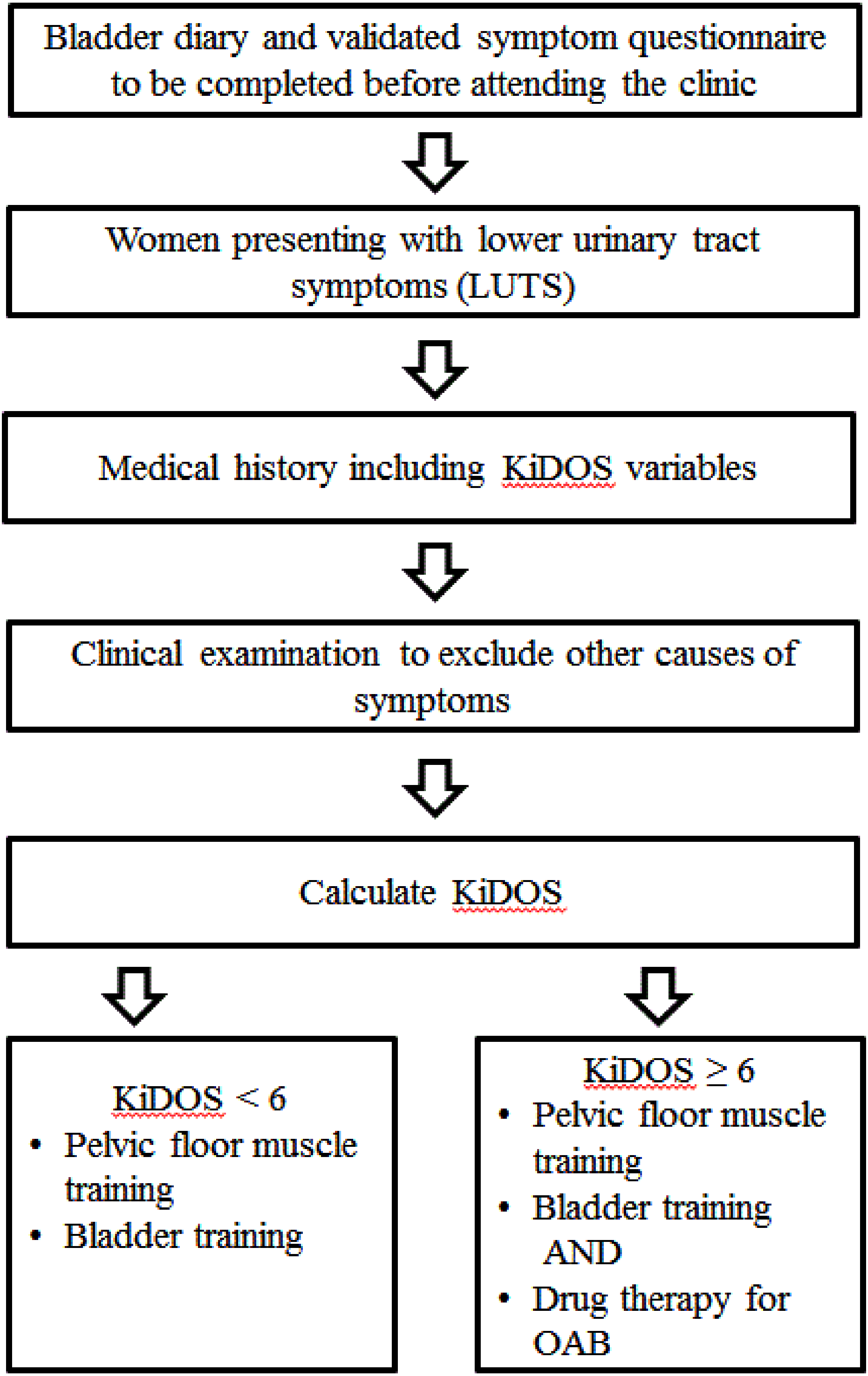
ROC: receiver operating characteristic

Table 6 Sensitivity, specificity, positive and negative predictive value for urgency and urgency incontinence

	Urgency	Urgency urinary incontinence
Sensitivity	96.2%	88.2%
Specificity	28.1%	47.6%
Positive predictive value (PPV)	44.2%	49.9%
Negative predictive value (NPV)	92.7%	87.2%



1.00
0.75
0.50
0.25
0.00
Sensitivity
0.00
0.25
0.50
0.75
1.00
1 - Specificity
Area under ROC curve = 0.7479


Bladder diary and validated symptom questionnaire to be completed before attending the clinic
Women presenting with lower urinary tract symptoms (LUTS)
Medical history including KiDOS variables
Clinical examination to exclude other causes of symptoms
Calculate KiDOS
KiDOS < 6
• Pelvic floor muscle training
• Bladder training
KiDOS ≥ 6
• Pelvic floor muscle training
• Bladder training
AND
• Drug therapy for OAB